Amostras emparelhadas

Professor: Pedro M. Almeida-Junior

21 de maio de 2021

Departamento de Estatística (UEPB)

Testes duas amostras dependentes

- Testes estatísticos de uma amostra envolvendo duas medidas ou replicações emparelhadas são usados quando o pesquisador deseja estabelecer se dois tratamentos são diferentes ou se um tratamento é melhor do que outro.

- O tratamento pode ser qualquer ocorrência de uma ampla variedade de condições:
    - Injeção de uma droga;
    - Métodos de ensino;
    - Treinamento de um operador na indústria;
    - Alteração cirúrgica;
    - Técnicas vocais

- Em cada caso, o grupo que tenha recebido o tratamento é comparado com um que não o tenha recebido ou que tenha recebido um outro tratamento.

- Em tais comparações de dois grupos, algumas vezes são observadas diferenças significantes as quais não são resultantes do tratamento.

- A ideia aqui é relacionar duas amostras estudadas, usando cada sujeito como seu próprio controle ou então formando pares de sujeitos e atribuindo as duas condições aos dois membros de cada par

- O indivíduo é exposto a ambos os tratamentos em momentos diferentes.

- A técnica paramétrica usual para analisar dados de duas amostras relacionadas é aplicar o teste t aos escores-diferença.

- Um escore-diferença pode ser obtido a partir dos dois escores dos dois membros de cada par combinado ou de dois escores de cada sujeito submetido às duas condições.

- O teste t assume que os escores-diferença são extraídos independentemente de uma distribuição normal, o que implica que as variáveis são medidas pelo menos em uma escala intervalar.

- Algumas vezes o teste t não é apropriado. O pesquisador pode se deparar com fatos como:
    - As suposições e as exigências do teste t não são satisfeitas aos nossos dados.
    - Níveis de mensuração abaixo da intervalar.
    - Os escores são simplesmente classificatórios
    - Os dados não são distribuídos normalmente

Uso de testes não paramétricos

- Quando as suposições do teste t não são atendidadas, o investigador pode escolher um dos **testes estatísticos não paramétricos para duas medidas de uma amostra**, ou replicações emparelhadas, os quais são apresentados nesta aula.

Teste de McNemar

- O **teste de McNemar** para a significância de mudanças é particularmente aplicável a modelos "**antes e depois**", em que cada sujeito é usado como seu próprio controle e nos quais as medidas são feitas em uma escala nominal ou ordinal.

- Pode ser usado para testar a eficácia de um tratamento particular (reunião, editorial de jornal, visita pessoal, campanha verbal, etc.) sobre as preferências de eleitores entre candidatos para um cargo eletivo.

- Note que esses são estudos nos quais as pessoas podem servir como seus próprios controles e nos quais medidas nominais ou categóricas seriam apropriadas para determinar a mudança "antes e depois".

O método

- Para testar a significância de qualquer mudança observada por este método, é usada uma tabela de frequências de quatro partes para representar o primeiro e o segundo conjunto de respostas de alguns indivíduos.

- As características gerais de tal tabela estão ilustradas na Tabela seguinte, na qual $+$ e $-$ são usados para denotar respostas diferentes.

	Depois	
Antes	$-$	$+$
$+$	A	B
$-$	C	D

As entradas na tabela são as frequências de ocorrência dos resultados associados:

- A denota o número de indivíduos cujas respostas foram $+$ na primeira medida e $-$ na segunda medida;

- D é o número de indivíduos que mudaram de $-$ para $+$;

- B é a frequência de indivíduos que responderam a mesma $(+)$ em cada ocasião;

- C é o número de indivíduos que responderam a mesma $(-)$ antes e depois.

- $A + D$ é o número total de pessoas cujas respostas mudaram.

- A hipótese nula é que o número de mudanças em cada direção é igualmente provável.

- Dos $A + D$ indivíduos que mudaram, esperaríamos:
  - $(A + D)/2$ indivíduos mudando de $+$ para $-$
  - $(A + D)/2$ indivíduos mudando de $-$ para $+$.

- Em outras palavras, quando H_0 é verdadeira, a frequência esperada em cada uma das duas células é $(A + D)/2$.

- Relembrando que
$$X^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$
em que $O_i =$ número observado de casos na i-ésima categoria $E_i =$ número esperado de casos na i -ésima categoria quando H_0 é verdadeira $k =$ número de categorias;

- No teste de McNemar para a significância de mudanças, estamos interessados somente nas células em que podem ocorrer mudanças.

$$
\begin{aligned}
X^2 &= \sum_{i=1}^{2} \frac{(O_i - E_i)^2}{E_i} \\
&= \frac{[A-(A+D)/2]^2}{(A+D)/2} + \frac{[D-(A+D)/2]^2}{(A+D)/2}
\end{aligned}
$$

Desenvolvendo a equação acima, temos

$$
X^2 = \frac{(A-D)^2}{A+D} \quad \text{com } gl = 1
$$

A distribuição amostral de X^2 quando H_0 é verdadeira é distribuída assintoticamente como um qui-quadrado com $gl = 1$.

CORREÇÃO PARA CONTINUIDADE

- A aproximação pela distribuição qui-quadrado da distribuição amostral de X^2 torna-se mais precisa se uma correção para continuidade for feita.

- A correção é necessária porque uma distribuição contínua (qui-quadrado) é usada para aproximar uma distribuição discreta $\left(X^2\right)$.

- Quando todas as frequências esperadas são pequenas, a aproximação pode ser pobre. O objetivo da correção para continuidade **(Yates, 1934)** é remover esta fonte de imprecisão.

CORREÇÃO PARA CONTINUIDADE

- Com a correção para continuidade incluída,

$$X^2 = \frac{(|A - D| - 1)^2}{A + D} \quad \text{com } gl = 1 \tag{1}$$

- A expressão do numerador na Equação (1) manda subtrair 1 do valor absoluto da diferença entre A e D (isto é, a diferença entre A e D independente do sinal) antes de elevar ao quadrado.

- A significância de qualquer valor observado de X^2, calculado a partir da Equação (1), é determinada da distribuição amostral χ^2 com $gl =$ 1.

Exemplo: Durante campanhas presidenciais aconteceram debates pela televisão entre dois ou mais candidatos. Um pesquisador estava interessado em determinar se um particular debate entre dois candidatos na eleição presidencial de 1980 foi eficaz em mudar as preferências dos telespectadores pelos candidatos. Para estabelecer a eficácia do debate, o pesquisador selecionou 75 adultos aleatoriamente antes do debate e pediu a eles para indicarem suas preferências pelos dois candidatos. Após a conclusão do debate, ele perguntou às mesmas pessoas por suas preferências pelos dois candidatos. Assim, em cada caso, ele sabia a preferência de cada pessoa antes e depois do debate.

Resultado da pesquisa

	Preferência após debate na TV	
Preferência antes debate na TV	Reagan	Carter
Carter	13	28
Reagan	27	7

Hipóteses do teste

- **Hipótese nula** H_0 : entre aqueles espectadores que mudaram suas preferências, a probabilidade de que um espectador troque de Reagan para Carter será igual à probabilidade de que o espectador troque de Carter para Reagan.

- A **hipótese alternativa** é H_1 : existe uma mudança diferenciada na preferência. As hipóteses podem ser resumidas como seguem:

$$H_0 : P[\text{ Reagan } \rightarrow \text{ Carter }] = P[\text{ Carter } \rightarrow \text{ Reagan }]$$
$$H_1 : P[\text{ Reagan } \rightarrow \text{ Carter }] \neq P[\text{ Carter } \rightarrow \text{ Reagan }]$$

Estatística do teste

- Para estes dados,

$$\begin{aligned} X^2 &= \frac{(|A - D| - 1)^2}{A + D} \quad \text{com } gl = 1 \\ &= \frac{(|13 - 7| - 1)^2}{13 + 7} \\ &= 5^2/20 \\ &= 1,25 \end{aligned}$$

Decisão

- Quando H_0 é verdadeira e $gl = 1$, a probabilidade de que $X^2 \geq 3,84$ é 0,05.

- **Decisão:** Como o valor observado de $X^2(1,25)$ é menor do que o valor crítico da quiquadrado (3,84), **não podemos rejeitar a hipótese** de que os candidatos eram igualmente eficientes em mudar as preferências dos espectadores.

FREQUÊNCIAS ESPERADAS PEQUENAS

- Foi observado anteriormente que a distribuição amostral de X^2 no teste qui-quadrado (e, portanto, no teste de mudança de McNemar) é bem aproximada pela distribuição qui-quadrado somente quando o tamanho da amostra é grande.

- Para amostras pequenas, a aproximação é pobre. No entanto, existe um procedimento alternativo quando N é pequeno.

- Se a frequência esperada para o teste de McNemar, $(A + D)/2$, é muito pequena (menor do que 5), deve ser usado o teste binomial no lugar do teste de McNemar. Para usar o teste binomial, considere $N = A + D$ e x a menor entre as duas frequências observadas, A ou D;

- Deve-se ressaltar que poderíamos ter analisado os dados do último exemplo usando o teste binomial (**checar como exercício**).

- Neste caso, a hipótese nula seria que a amostra de $N = A + D$ casos veio de uma população binomial onde $p = q = \frac{1}{2}$. Para os dados acima, $N = 20$ e $x = 7$, a menor entre as duas frequências observadas.

Passos no cálculo do teste de mudança de McNemar

- Coloque as frequências observadas em uma tabela de contingência 2×2.
- Determine o número total de "mudanças", $A + D$. Se o número total de mudanças é **menor do que 10** , use o teste binomial em vez do teste de McNemar.
- Se a frequência total de mudanças excede 10 , calcule o valor de X^2 usando a Equação (1).
- Determine a probabilidade associada com um valor tão grande quanto o valor observado de X^2 obtido pela qui-quadrado.
- Se a probabilidade para o valor observado de X^2 com $gl = 1$ é menor ou igual a α, rejeite H_0 em favor de H_1.

Exercício

Exercício: Foram avaliados 100 dentes molares com o objetivo de detectar cáries. Os mesmos 100 dentes foram avaliados de duas formas diferentes: com Raio X e visualmente obtendo-se os seguintes resultados:

	Raio X	
Visualmente	Com cáries	sem cáries
Sem cáries	4	45
Com cáries	34	17

Testar, usando $\alpha = 0,05$, se há diferença em usar os dois métodos para detectar cáries.

Exercício: Foram avaliados 100 dentes molares com o objetivo de detectar cáries. Os mesmos 100 dentes foram avaliados de duas formas diferentes: com Raio X e visualmente obtendo-se os seguintes resultados:

	Raio X	
Visualmente	Com cáries	sem cáries
Sem cáries	4	45
Com cáries	34	17

Testar, usando $\alpha = 0,05$, se há diferença em usar os dois métodos para detectar cáries.

Resposta: H_0 : A proporção de dentes molares com cáries usando Raio X é igual a proporção de dentes molares usando o método visual. H_1 : As proporções envolvidas são diferentes. $\chi^2 = 6,857 > \chi^2_{(1,0,05)} = 3,84 \Rightarrow$ rejeitar H_0